

多特征和多视角信息融合的步态识别

赵永伟 张二虎 鲁继文 胡峻林

(西安理工大学信息科学系, 西安 710048)

摘要 提出了一种基于多特征和多视角信息融合的步态识别方法。应用背景差分 and 阴影消除获得人体步态轮廓, 对人体轮廓使用伪 Zernike 矩、小波描述子和 Procrustes 形状分析法进行了特征提取。通过多特征和多视角步态信息融合, 完成了基于人体步态特征的身份识别。该方法在 CASIA 步态数据库上进行了实验, 取得了较高的正确识别率, 实验结果表明本文所提出的识别方法具有较高的识别性能。

关键词 步态识别 伪 Zernike 矩 小波描述子 Procrustes 形状分析 信息融合

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2009)03-0388-06

Gait Recognition Via Multiple Features and Views Information Fusion

ZHAO Yong-wei, ZHANG Er-hu, LU Ji-wen, HU Jun-lin

(Department of Information Science, Xi'an University of Technology, Xi'an 710048)

Abstract A new gait recognition method based on information fusion of multiple kinds of features and views is proposed in this paper. Through the background subtraction and shadow elimination, human motion silhouettes are obtained and gait features are extracted using pseudo-Zernike moment, wavelet descriptor and Procrustes shape analysis. The gait recognition is accomplished through information fusion of multiple kinds of features and views on feature level and decision level. The method is evaluated on the CASIA gait database and received comparative high correct recognition rate. The experimental results show that our approach has efficient recognition performance.

Keywords gait recognition, pseudo-Zernike moment, wavelet descriptor, Procrustes shape analysis, information fusion

1 引言

近年来,随着一些安全敏感场合(如机场、军事基地等)对智能监控系统的迫切需求,非接触式远距离的身份识别技术备受关注。但由于人脸、指纹、虹膜等生物特征识别技术都需要人体近距离或接触性的感知,不适合于远距离下的身份识别。而进入监控范围的人,其运动步态是可见的,并且具有易采集、非侵犯性、难以伪装等优点,近年来已引起了国内外智能视觉监控研究者的浓厚兴趣。

步态识别就是根据人走路的姿势进行身份识

别。近年来,国内外学者对步态识别进行了深入地研究,相继取得了一些成果,但大都只停留在算法的研究上。目前文献报道的步态识别方法大致可以分为两类:基于模型和基于统计的方法。基于模型的方法重点关注人的运动信息,预先建立模型,通过模型和图像序列的匹配获得模型参数,使用这些参数作为步态特征进行分类。如 Cunado 等人^[1]将人体下半身看作一个骨架,建立骨架的运动模型,再用速度霍夫变换(VHT)提取运动参数;Lee^[2]将人的身体侧面二值化图像的不同部分用7个椭圆拟合,每一个椭圆用质心等4个特征表示,加上整个身体图像质心的高度共29个特征表示整个人体侧面图像。

基金项目: 陕西省自然科学基金研究计划项目(2006F26); 陕西省教育厅自然科学研究项目(07JK353)

收稿日期: 2007-09-18; 改回日期: 2007-12-12

第一作者简介: 赵永伟(1984~),男。西安理工大学印刷包装工程学院信号与信息处理专业硕士研究生。主要研究方向为数字图像处理与模式识别。E-mail: waboy_zhao@163.com

基于统计的方法重点关注人运动的静态信息,从图像序列中提取步态的统计参数。如王亮等人^[3-4]提出的基于轮廓的步态识别方法,Shutler等人^[5]提出的基于时间矩的步态识别方法都是统计识别方法。

不同于上述方法,本文提出基于多特征和多视角信息融合的步态识别方法,首先通过背景减除技术获得人体运动轮廓,并对其应用伪 Zernike 矩、小波描述子和 Procrustes 形状分析法进行特征提取,然后在各个视角下进行多特征融合,最后利用不同视角下的步态信息互补这一特点,完成了多视角下步态信息融合的身份识别。算法流程如图 1 所示,并在 CASIA 步态数据库上进行了实验,给出了实验结果。



图2 人体轮廓提取结果

Fig. 2 The result of edge extraction

目前,使用的步态特征主要有基于运动和基于统计两种^[6-8]。为了保留步态序列中的动态及静态信息,本文使用伪 Zernike 矩及小波描述子来描述步态序列中每一帧的人体轮廓,而使用 Procrustes 形状分析法来对人体步态运动图像集进行统计描述,从而有效地克服了单一步态特征唯一性不强、稳定性不够等缺陷。

2.1 伪 Zernike 矩

伪 Zernike 矩是基于图像整个区域的形状描述算子^[9-10]。由于伪 Zernike 矩的基是正交径向多项式,与其他矩算子相比,伪 Zernike 矩描述子具有较强的旋转、平移和比例不变性,信息表达的高效性以及能较好地描述运动图像等特点,因此采用它来描述人体步态特征。

伪 Zernike 矩是一种正交复数矩,它所利用的正交多项式集是一个在单位圆($D^2: x^2 + y^2 \leq 1$)内的完备正交集。对于数字图像,设 $I(x, y)$ 是当前像素,则 p 阶 q 次伪 Zernike 矩 P_{pq} 可以定义为

$$p_{pq} = \frac{p+1}{\pi} \sum_x \sum_y I(x, y) W_{pq}^*(x, y) \quad (1)$$

式中, p 是正整数或零, q 是整数,并且 $|q| \leq p$, $*$ 表示复数共轭,并且 $x^2 + y^2 \leq 1$ 。 $W_{pq}(x, y)$ 是 p 阶 q 重的伪 Zernike 矩基函数,在极坐标中 p 阶 q 重的伪 Zernike 矩基函数可表示为

$$W_{pq}(\rho \cos \theta, \rho \sin \theta) = S_{pq}(\rho) \exp(iq\theta) \quad (2)$$

式中, ρ, θ 分别表示 (x, y) 在极坐标下的模和幅角, i 是虚数单位, S_{pq} 是由下式给出的径向多项式:

$$S_{pq}(\rho) = \sum_{s=0}^{p-q} \frac{(-1)^s (2p-1-s)!}{s!(p-q-s)!(p+q+1-s)!} \rho^{p-s} \quad (3)$$

本文使用了 0 到 9 阶伪 Zernike 矩,共 55 个特征值来描述一个轮廓图像,如图 3 所示。

2.2 小波描述子

小波描述子是一种可以很好地对 2 维形状的轮廓进行描述的数学工具,它最大的优点就是对于形

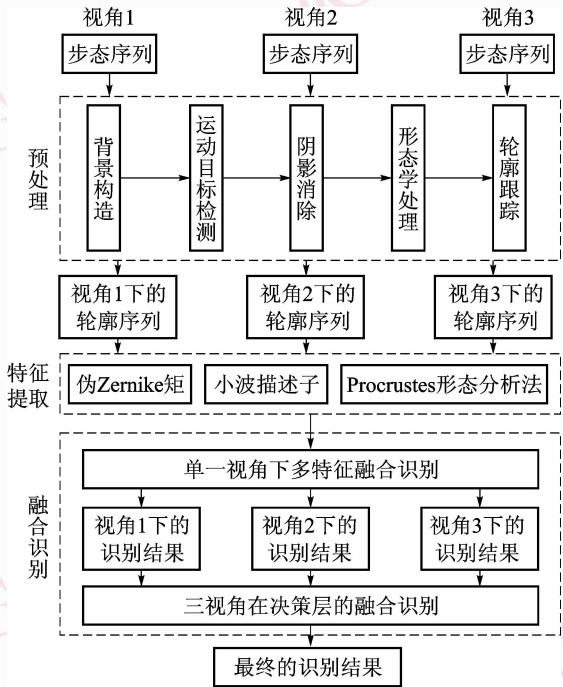


图1 基于多特征和多视角信息融合的步态识别流程图
Fig. 1 Flow chart of the gait recognition based on multiple features and views information fusion

2 特征提取

基于人体步态特征的身份识别很大程度上依赖于人体轮廓形状随着时间变化的规律,通过背景减除、二值化及阴影消除后,可以得到步态序列中每帧图像的单连通人体运动区域。通过轮廓跟踪的方法,可以快速得到人体轮廓,图 2 给出了其中一帧的轮廓提取结果。

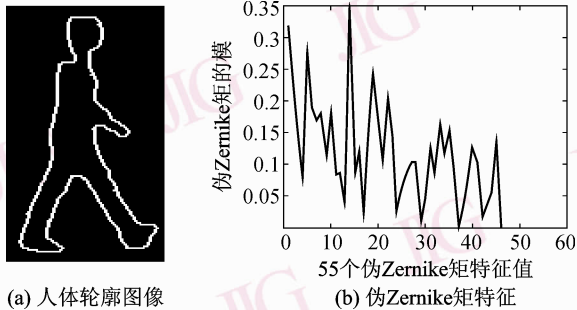


图 3 伪 Zernike 矩提取人体轮廓特征

Fig. 3 Pseudo-Zernike feature of a silhouette

状的旋转、尺度及平移有很高的鲁棒性^[11]。使用小波描述子来表示一个形状,可以同时得到该形状的时域及频域信息。

本文使用离散小波变换来描述人体的轮廓。对于步态图像序列中的第 i 帧轮廓图像,选定头顶边缘点作为参考起点,沿逆时针方向在轮廓边界上选取 k 个点,计算每个点到轮廓质心的欧氏距离,则该轮廓可以表示为一个由 k 个元素组成的向量 $\mathbf{D}_i = [d_{i1}, d_{i2}, \dots, d_{ik}]$ 。使用边界像素的内插处理来解决点的匹配问题,以便对于每幅图像而言点数是相同的,本文使用 $k = 128$ 。另外,为了消除图像尺度对于训练和识别过程的影响,还必须对距离进行归一化处理。

$\mathbf{D}_i$ 可以看作是一个 1 维信号,选择小波基 h 后,使用式(4)对 $\mathbf{D}_i$ 进行两层小波变换,得到第 i 帧轮廓图像的小波描述子 $\mathbf{W}_i$ 。

$$\mathbf{W}_i = \langle \langle \mathbf{D}_i, h \rangle, h \rangle \quad (4)$$

为了压缩数据及计算方便,每帧图像只使用低频段的 32 个点来进行匹配识别。用小波描述子来提取人体轮廓特征的结果如图 4 所示。

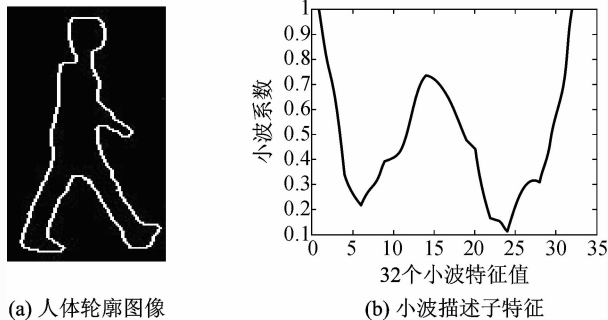


图 4 小波描述子提取人体轮廓特征

Fig. 4 Wavelet feature of a silhouette

2.3 Procrustes 形状分析法

统计识别方法通常对运动图像集进行统计描述。Procrustes 形状分析法是方向统计学中一种特别流行的方法,它非常适用于编码 2 维形状,并且提供了一种好的工具来寻找一组形状的均值形状^[12],因此可采用它来描述人体步态特征。假设 2 维空间中一个形状可以被描述为一个由 k 个复数组成的向量,或称作一个配置,即 $\mathbf{Z} = [z_1, z_2, \dots, z_k]^T$ 。通过定义中心配置,即 $\mathbf{U} = [u_1, u_2, \dots, u_k]^T$, $u_i = z_i - \bar{z}$, $\bar{z} = \sum_{i=1}^k \frac{z_i}{k}$, 可以将每个形状置于空间坐标的中心。两个配置 $\mathbf{U}_1, \mathbf{U}_2$ 之间的 Procrustes 距离被定义为:

$$d(\mathbf{U}_1, \mathbf{U}_2) = 1 - \frac{|\langle \mathbf{U}_1^*, \mathbf{U}_2 \rangle|^2}{\|\mathbf{U}_1\|^2 \|\mathbf{U}_2\|^2} \quad (5)$$

式中, $0 \leq d(\mathbf{U}_1, \mathbf{U}_2) \leq 1$ 。

Procrustes 距离可以在独立于位置、尺度和旋转的情况下比较两个形状的相似性,Procrustes 距离越小,则两个形状越相似。

给定一个包含 n 个形状的集合,可以通过寻找 $\bar{\mathbf{U}}$ 来获得它们的均值形状。为了计算 $\bar{\mathbf{U}}$,先计算下面的配置矩阵:

$$\mathbf{S} = \sum_{i=1}^n (\langle \mathbf{U}_i \mathbf{U}_i^* \rangle / \langle \mathbf{U}_i^* \mathbf{U}_i \rangle) \quad (6)$$

则 Procrustes 均值形状 $\bar{\mathbf{U}}$ 对应矩阵 $\mathbf{S}$ 的最优支配向量,即对应配置矩阵 $\mathbf{S}$ 最大特征值的特征向量。

采用 Procrustes 形状分析法提取步态特征的步骤如下:

(1) 从人体轮廓边界上选择 k 个点来表示人体轮廓形状为一个向量配置 $\mathbf{Z}_j$ ($j = 1, 2, \dots, n$, n 是一个步态图像序列中的帧数),本文使用 $k = 128$;

(2) 设置中心配置。由于坐标中心是人体的质心,所以可以直接设置中心配置 $\mathbf{U}_j = \mathbf{Z}_j$;

(3) 使用式(6)计算配置矩阵 $\mathbf{S}$,然后计算该矩阵的特征值和对应的特征向量;

(4) 设置 Procrustes 均值形状为相应于 $\mathbf{S}$ 的最大特征值的特征向量,这个均值形状将作为步态序列的统计静态特征用于识别。

3 个视角下同一个人的不同步态序列的 Procrustes 均值形状及不同人的步态序列的 Procrustes 均值形状如图 5 和图 6 所示。可见,同一个人的不同步态序列的均值形状都非常接近,而不同人的步态序列的

均值形状相差就比较大。

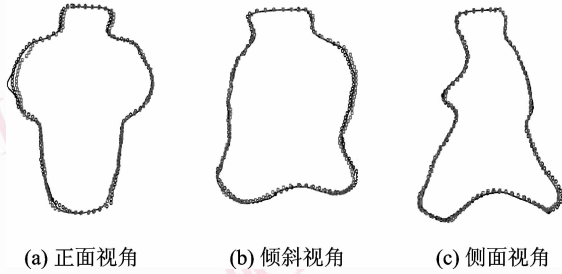


图5 3个视角下同一个人的不同步态序列的均值形状

Fig.5 Procrustes average shapes of a person's different gait sequences in three views

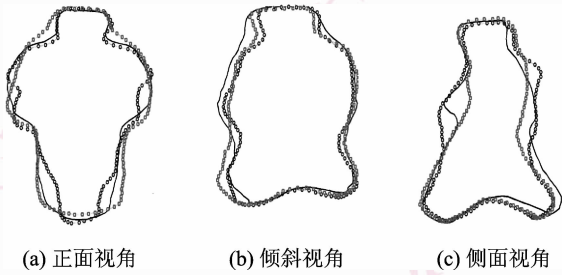


图6 3个视角下不同人的步态序列的均值形状

Fig.6 Procrustes average shapes of three person's gait sequences in three views

3 识别

3.1 单视角下的多特征融合

考虑到单一特征往往存在着唯一性不强与鲁棒性不够等缺陷,本文在各个视角下进行了3种特征融合识别,以提高正确识别率。本文使用计分^[13]的方法进行多特征的融合识别,计分方法如下

$$Score(X, c) = \sum_{i=1}^3 v_i^j \mu_i(X, c) \quad (7)$$

上式中, X 是待识别步态序列, c 是训练步态序列。 v_i^j 是一个权值,其值来自第 i 种特征在 j 视角下的正确识别率,是一个先验概率。 $\mu_i(X, c)$ 表示待识别步态序列 X 对训练步态序列 c 的隶属度,可以通过下式计算获得

$$\mu_i(X, c) = \frac{1 / \|X - c\|^{2/(b-1)}}{\sum_{j=1}^k 1 / \|X - X_j\|^{2/(b-1)}} \quad (8)$$

式中, $b > 1$ 是一个可以控制聚类结果的模糊程度的常数。

3.2 多视角步态信息融合识别

由于不同视角下的人体步态存在着较强的信息互补的特点,又分别采用了基于 Rank、基于 KNN (k-nearest neighbor) 以及基于 Fuzzy 等融合机制^[14-20],完成了对多视角步态信息的融合识别。

(1) 基于 Rank 的融合决策

基于 Rank 的融合决策就是将每个人在 3 个视角下的正确识别时的 Rank 和计算出来,具有最小的 Rank 和的类就为组合分类器的融合结果。令 $r(C, R)$ 为等级 R 拥有类别 C 的等级 Rank,则组合计算方法为

$$R_F = \left\{ \left(C, \sum_{i=1}^m r(C, R) \right), C \in S \right\} \quad (9)$$

式中, S 表示步态样本集。

(2) 基于 KNN 的融合决策

基于 KNN 的融合决策就是在 Rank = 1 的情况下,在训练样本中找出原始 n 个模式识别模块输出。对于待识别的步态序列 X ,设 $k_1, k_2, \dots, k_n$ 分别代表输出中属于第 1 类到第 n 类的样本个数,则融合规则判别函数可定义为

$$\text{if } g_c(X) = \max_j k_j \text{ then } X \in c \quad (10)$$

(3) 基于 Fuzzy 的融合决策

基于 Fuzzy 的融合决策先使用式(8)计算待识别序列对样本库中的所有样本序列的隶属度,取 3 个视角下隶属度之和最大的一个就是最终的识别结果。其融合规则判别函数可定义为

$$\text{if } g_c(X) = \max_j \sum_{i=1}^3 v_i^j \mu_{ij} \text{ then } X \in c \quad (11)$$

3.3 实验结果

应用上述方法在中科院自动化所的 CASIA 步态数据库^[21]上进行实验。CASIA 步态数据库包含 124 个人的步态序列,分别使用 11 个摄像机从 11 个不同的视角同时进行拍摄。这 11 个视角分别是 0, 18, 36, ..., 180,而每一个视角下都有 10 个步态序列,分别是 2 个提包走路的序列、2 个穿着大衣走路的序列以及 6 个在普通条件下走路的序列。CASIA 步态数据库中总共有 $124 \times 11 \times 10 = 13\ 640$ 个步态序列,所有的步态序列图像的分辨率都是 320×240 像素。

CASIA 步态数据库把视角、衣着和附加物等各种影响步态识别的因素都考虑了,而在实验中,由于只考虑不同视角下的信息融合,所以只使用在普通条件下走路的 6 个序列进行测试。另外,考虑到两个比较接近的视角的信息冗余度比较大,所以在实

验中,只选取正面视角(0°)、倾斜视角(54°)和侧面视角(90°)进行测试,然后在这 3 个视角的基础上进行多特征和多视角融合识别。我们采用交叉验证进行实验,先把每个人的 6 个序列分成 3 组,任取其中 2 组进行训练,剩下 1 组进行识别,总共进行了 $C_3^6=3$ 次实验。表 1 给出了不同视角下各种特征在 NN(nearest neighbor)和 KNN 分类器下的平均正确识别率,表 2 是各个视角下多种特征进行融合后的总的识别结果。

表 1 各个视角下采用不同步态特征的识别结果

Tab. 1 Recognition result of three kinds of features in three views

步态特征	分类器	正面视角 (%)	倾斜视角 (%)	侧面视角 (%)
伪 Zernike 矩	NN	91.13	87.90	91.13
	KNN	84.68	89.52	89.52
小波描述子	NN	74.19	69.35	72.58
	KNN	70.97	67.74	75.81
Procrustes	NN	92.34	80.65	87.10
	KNN	87.10	82.26	88.71

表 2 各个视角下的多特征融合识别结果

Tab. 2 Fusion result of multiple kinds of features in three views 单位:%

分类器	正面视角	倾斜视角	侧面视角
NN	94.35	90.32	93.55
KNN	95.16	90.32	95.16

得到单一视角下的识别结果后,采取基于 Rank、基于 KNN 和基于 Fuzzy 的决策层融合规则进行多视角信息融合识别。同时也使用了 Yuan Wang 等人在文献[21]中提出的 4 种融合机制(Sum Rule, Weighted Sum Rule, Product Rule, Dempster-Shafer Rule)进行了测试,表 3 给出了使用以上 7 种融合机制的融合识别结果。

表 3 3 个视角的融合识别结果

Tab. 3 Recognition result of multiple views fusion 单位:%

文献[20]中的融合方法				本文的融合方法		
Sum	W-Sum	Product	D-S	Rank	KNN	Fuzzy
95.56	96.37	95.97	97.17	96.77	95.56	99.19

根据以上的识别结果,可以得出以下几点结论:

- (1)在正面视角下,Procrustes 形状分析法的识别率比伪 Zernike 矩及小波描述子的识别率都高;在倾斜和侧面视角下,伪 Zernike 矩的识别率比 Procrustes 形状分析法的识别率偏高。这是因为正面视角下的步态序列的动态特征不明显,无法得到手和脚的摆动角度等动态信息,是正面人体由远到近的过程,相邻帧之间的人体形状基本上只有大小上的变化,在形状上变化不大,所以适合使用 Procrustes 形状分析法进行统计描述;而在倾斜和侧面视角下的步态序列的动态特征就较为明显,相邻帧之间的人体形状具有较大的差异,对每一帧的人体形状进行描述可以得到帧与帧之间的变化,保留步态的动态特征;
- (2)在各个视角下进行多特征融合后,正面视角的正确识别率最高,倾斜视角最低。这表明,单一视角下的步态识别有一定的局限性;
- (3)在单一视角下基于多特征融合的步态识别比任一特征在单一视角下的识别率都高;
- (4)基于多视角信息融合的步态识别比任一视角下的识别率都高;
- (5)提出的基于 Fuzzy 的融合机制的融合效果最好,其次是 Dempster-Shafer 融合规则。因为它们都考虑了待识别序列和参考序列之间的距离,而基于 Fuzzy 的融合机制还考虑了各个视角的可信度,即式(11)中的权值 v_i^j ,所以获得了更高的正确识别率。

4 结 论

提出了一种基于多特征和多视角信息融合的步态识别方法。首先使用伪 Zernike 矩、小波描述子和 Procrustes 形状分析法对人体步态特征进行了提取与变换;然后在各个视角下进行多特征融合,并应用近邻分类器进行步态的分类与识别;最后,通过融合不同视角下的人体步态信息,完成了多视角下步态信息融合的身份识别。实验表明该方法具有较高的识别性能,是一种有效的步态识别算法。未来的工作重点将集中于建立更准确地反映人体运动的模型,解决人携带物体行走的问题,以及利用 3 个视角的信息进行 3D 重构从而解决视角选择的问题。

致 谢 本文使用了中科院自动化所的 CASIA 步态数据库进行实验,于此表示感谢。

参考文献 (References)

- 1 Cunado D, Nixon M S, Carter J N. Automatic gait recognition via model-based evidence gathering [A]. In: Proceedings of IEEE Workshop on Automated Identification Advanced Technologies[C], Summit, New Jersey, USA, 1999: 27-30.
- 2 Lee Lily. Gait Analysis for Classification[R]. AI Technical Report 2003-014, Cambridge, Massachusetts, USA; Massachusetts Institute of Technology, 2003.
- 3 Wang Liang, Tan Tie-niu. Automatic gait recognition based on statistical shape analysis [J]. IEEE Transactions on Image Processing, 2003, **12**(9): 1120-1131.
- 4 Wang Liang, Tan Tie-niu, Ning Hua-zhong. Silhouette analysis based gait recognition for human identification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, **25**(12): 1505-1518.
- 5 Shutler J, Nixon M, Harris C. Statistical gait recognition via temporal moments[A]. In: Proceedings of IEEE Southwest Symposium on Image Analysis and Interpretation [C], Austin, Texas, USA, 2000: 291-295.
- 6 Wang Liang, Hu Wei-ming, Tan Tie-niu. Gait-based human identification[J]. Chinese Journal of Computers, 2003, **26**(3): 353-360. [王亮, 胡卫明, 谭铁牛. 基于步态的身份识别[J]. 计算机学报, 2003, **26**(3): 353-360.]
- 7 Cunado D, Nixon M S, Carter J N. Automatic extraction and description of human gait models for recognition purposes [J]. Computer Vision and Image Understanding, 2003, **90**(1): 1-41.
- 8 Su Kai-na, Liu Yu-dong, Ma Li. The extraction of gait features in human identification [J]. Journal of Beijing University of Technology, 2005, **31**(4): 388-393. [苏开娜, 刘玉栋, 马丽. 身份识别中步态特征的提取[J]. 北京工业大学学报, 2005, **31**(4): 388-393.]
- 9 Shutler J D, Nixon M S. Zernike velocity moments for sequences-based description of moving features [J]. Image and Vision Computing, 2006, **24**(4): 343-356.
- 10 Ye Bin, Peng Jia-xiong. Improvement and invariance analysis of pseudo-zernike moments[J]. Journal of Image and Graphics, 2003, **8**(3): 246-252. [叶斌, 彭嘉雄. 伪 Zernike 矩不变性分析及其改进研究[J]. 中国图象图形学报, 2003, **8**(3): 246-252.]
- 11 Zhao Guo-ying, Li Cui, Li Hua. Gait recognition using fractal scale and wavelet moments [A]. In: Proceedings of International Conference on Pattern Recognition[C], HongKong, 2006: 453-456.
- 12 Wang Liang, Ning Hua-zhong, Hu Wei-ming. Gait recognition based on procrustes shape analysis[A]. In: Proceedings of International Conference on Image Processing[C], Rochester, New York, USA, 2002: 22-25.
- 13 Boulgouris N V, Plataniotis K N, Hatzinakos D. Gait recognition using linear time normalization [J]. Pattern Recognition, 2006, **39**(5): 969-979.
- 14 Liu Zong-yi, Sudeep Sarkar. Improved gait recognition by gait dynamics normalization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, **28**(6): 863-876.
- 15 Lam T H, Lee R S T. Human identification by using the motion and static characteristic of gait[A]. In: Proceeding of 18th International Conference on Pattern Recognition[C], HongKong, 2006: 996-999.
- 16 Begga R, Kamruzzaman J. A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data[J]. Journal of Biomechanics, 2005, **38**(3): 401-408.
- 17 Kuncheva L I, Bezdek J C, Duin R P W. Decision templates for multiple classifier fusion: An experimental comparison[J]. Pattern Recognition, 2001, **34**(2): 299-314.
- 18 Wang Liang, Ning Hua-zhong, Tan Tie-niu. Fusion of static and dynamic body biometrics for gait recognition[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, **14**(2): 149-158.
- 19 Lu Ji-wen, Zhang Er-hu, Xue Yan-xue. Gait recognition based on independent component analysis and information fusion[J]. Pattern Recognition and Artificial Intelligence, 2007, **30**(3): 365-370. [鲁继文, 张二虎, 薛延学. 基于独立成分分析和信息融合的步态识别[J]. 模式识别与人工智能, 2007, **20**(3): 365-370.]
- 20 Lu Ji-wen, Zhang Er-hu. Gait recognition for human identification based on ICA and fuzzy SVM through multiple views fusion[J]. Pattern Recognition Letters, 2007, **28**(16): 2401-2411.
- 21 Wang Yuan, Yu Shi-qi, Wang Yun-hong, et al. Gait recognition based on fusion of multi-view gait sequences[A]. In: Proceedings of International Association of Pattern Recognition International Conference on Biometrics[C], HongKong, 2006: 605-611.